



NIT-287

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Serial. No.: 09/917,829 Confirmation No.: 1721
Applicant: N. NUKAGA et al.
Filed: July 31, 2001
Title: VOICE SYNTHESIZING METHOD AND VOICE SYNTHESIZER
PERFORMING THE SAME
TC/AU: 2654
Examiner: V. Chawan
Customer No.: 24956

SUBMISSION OF CERTIFIED PRIORITY DOCUMENT

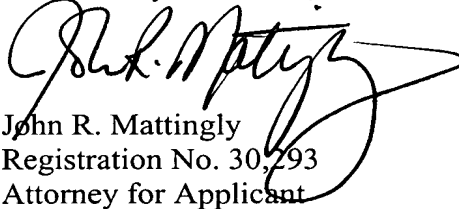
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Applicants submit herewith a certified priority document of corresponding Japanese Patent Application No. 2001-175090, filed June 11, 2001 for the purpose of claiming foreign priority under 35 U.S.C. § 119.

Applicants respectfully request that the priority document be submitted and filed to be officially considered of record.

Respectfully submitted,


John R. Mattingly
Registration No. 30,293
Attorney for Applicant

MATTINGLY, STANGER, MALUR & BRUNDIDGE, P.C.
1800 Diagonal Road, Suite 370
Alexandria, Virginia 22314
(703) 684-1120
Date: October 24, 2005

BEST AVAILABLE COPY

CERTIFIED COPY OF
PRIORITY DOCUMENT

日 本 国 特 許
JAPAN PATENT OFFICE

INVENTOR'S CERTIFICATE
OF PRIORITY DOCUMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office

出 願 年 月 日

Date of Application:

2001年 6月11日

出 願 番 号

Application Number:

特願2001-175090

出 願 人

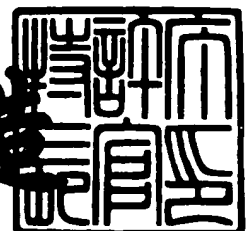
Applicant(s):

株式会社日立製作所

2001年 8月 3日

特 許 庁 長 官
Commissioner,
Japan Patent Office

及 川 耕 造



【書類名】 特許願

【整理番号】 NT01P0597

【提出日】 平成13年 6月11日

【あて先】 特許庁長官 殿

【国際特許分類】 G10L 3/00

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内

 【氏名】 額賀 信尾

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内

 【氏名】 永松 健司

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内

 【氏名】 北原 義典

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社日立製作所

【代理人】

 【識別番号】 100068504

 【弁理士】

 【氏名又は名称】 小川 勝男

 【電話番号】 03-3661-0071

【選任した代理人】

 【識別番号】 100086656

 【弁理士】

 【氏名又は名称】 田中 恭助

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100094352

【弁理士】

【氏名又は名称】 佐々木 孝

【電話番号】 03-3661-0071

【手数料の表示】

【予納台帳番号】 081423

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声合成方法及びそれを実施する音声合成装置

【特許請求の範囲】

【請求項 1】

定型的文を音声合成により音声にする音声合成方法であって、上記定型的文の発声内容の種別を指定する発声内容識別子定め、上記発声内容識別子に対応する話調及び韻律データからなる話調辞書を作成し、発生すべき合成音声の内容識別子及び話調を指定して、上記話調辞書から上記発生すべき合成音声の韻律データを選択し、上記選択された韻律データを音声合成装置駆動データとして音声合成手段に加え、特定の話調の音声合成を行うことを特徴とする音声合成方法。

【請求項 2】

上記韻律データデータは、少なくとも、上記定型的文の発声内容を発音音素に分解した発音記号列と、上記発音記号列を構成する各音素の長さ、高さ、強さの情報から構成されるデータであることを特徴とする請求項 1 記載の音声合成方法。

【請求項 3】

定型的文を韻律データに変換し、韻律データを音声合成装置駆動データとして音声合成合成処理部にくわえ音声合成を行う音声合成装置であって、上記定型的文の種別を指定する発声内容識別子と、合成音声の話調を指定する話調指定情報と、韻律データとが対応つけられた話調辞書を格納するメモリと、音声合成時に合成すべき音声の発声内容識別子と話調とを指定する手段と、上記指定する手段によって指定された韻律データを上記話調辞書から選択し、音声信号に変関する声合成合成処理部とを具備することを特徴とする音声合成装置。

【請求項 4】

上記韻律データデータは、少なくとも、上記定型的文の発声内容を発音音素に分解した発音記号列と、上記発音記号列を構成する各音素の長さ、高さ、強さの情報から構成されるデータであることを特徴とする請求項 3 記載の音声合成装置。

【請求項 5】

請求項 3 又は 4 に記載の音声合成装置をもつことを特徴とする携帯電話機。

【請求項 6】

定型的文を韻律データに変換し、韻律データを音声合成装置駆動データとして端末装置の音声合成部に加え音声合成を行う韻律データの配信方法であって、上記定型的文の発声内容の種別を指定する発声内容識別子定め、上記内容識別子に対応する話調及び韻律データからなる話調辞書を作成し、上記話調辞書を通信ネットワークに設けられたサーバ又は上記サーバを介して接続された端末装置に供給する音声合成を行う韻律データの配信方法。

【請求項 7】

上記韻律データは、少なくとも、上記定型的文の発声内容を発音音素に分解した発音記号列と、上記発音記号列を構成する各音素の長さ、高さ、強さの情報から構成されるデータであることを特徴とする請求項 6 記載の韻律データの配信方法。

【請求項 8】

請求項 6 又は 7 に記載の韻律データの配信方法において、上記話調辞書を通信ネットワークに設けられた上記サーバを介して接続された端末装置に供給する場合、上記端末装置は端末使用者が指定した話調に対応する話調辞書を指定する手段と、指定された話調辞書を上記サーバから上記端末装置に転送するデータ転送手段と、転送された話調辞書を、端末装置内の話調辞書格納メモリに格納する話調辞書格納手段を備えることにより、端末使用者が指定する話調で合成音を行うことを特徴とする韻律データの配信方法。

【請求項 9】

上記話調辞書の作成は公開された発声内容管理リストを参照することにより、韻律データを作成することを特徴とする請求項 7 又は 8 に記載の韻律データの配信方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声合成方法及びそれを実施する音声合成装置及びシステム、更に

詳しく言えば、音声合成すべき内容が略定まっている定型的文を音声に変換する
する音声合成方法、その方法を実施する音声合成装置及びその方法及び装置を実
施するに必要なデータの作成方法に関する。特に、音声合成装置をもつ携帯端末
及びそれと接続可能なデータ通信手段からなる通信網で利用される。

【0002】

【従来の技術】

一般に、音声合成は、発音する内容を示す発音記号（音素記号）と、音声の抑
揚の物理的尺度であるピッチの時系列パターン（基本周波数パターン）、及び各
音素の長さ（音素継続長）、及び強さ（音素強度）から、音声波形を生成する技
術である。以下、基本周波数パターン、音素継続長及び音素強度の三つのパラメ
ータを「韻律パラメータ」と総称し、音素記号と韻律パラメータとの組を「韻律
データ」と総称する。

【0003】

音声波形を生成する方式としては、音素の声道特性を模擬するパラメータをフ
ィルタで駆動するパラメータ合成方式と、人間の発声した音声波形から音素特徴
を示す断片を切り出して接続することにより波形を生成する波形接続方式が代表
的である。このように、音声合成においては、「韻律データ」を生成することが
重要である。また、上記音声合成方法は日本語のみならず、言語一般に共通して
用いることができる。

【0004】

音声合成では、合成対象となる文内容に対応する上記韻律パラメータを何らか
の方法で求める必要がある。例えば、電子メールや電子新聞の読み上げ等に音声
合成技術を適用する場合には、任意の文章を言語解析し、単語や文節の区切り位
置を同定し、文節のアクセント型を決定した後、アクセント情報や音節情報等か
ら韻律パラメータを求める必要がある。これらの自動変換に関する基本方式は既
に確立されており、「隣接単語間の結合関係に着目したテキスト音声変換用形態
素解析処理」（日本音響学会誌51巻1号、1995、pp. 3-13）に開示
されている方法で実現できる。

【0005】

上記韻律パラメータのうち、音節（音素）継続時間長は、音節（音素）が置かれるコンテキストを始めとする種々の要因によって変化する。継続時間長に影響を与える要因としては、当該音節の種類のような調音上の制約、タイミング、単語の重要度、発話区分境界の明示、発話区分内のテンポ、全体のテンポ、構文意味内容等の言語的制約等がある。継続時間長制御においては、実際に観測される継続時間長データに対して、上記要因に関する影響度等を統計的に分析し、その結果得られる規則を利用する方式が一般的である。例えば、「規則による音声合成のための音韻時間長制御」（電子通信学会論文誌、1984/7、Vol. J67-A、No. 7）には、上記韻律パラメータの計算方法が記載されている。もちろん、韻律パラメータの計算方法はこの限りではない。

【0006】

上述の音声合成方法は、任意の文から韻律パラメータに変換する方法、すなわちテキスト音声合成方法に関するものであるが、一方、合成すべき内容が予め定まっている定型的な文に対する音声合成する場合の韻律パラメータ計算方法がある。定型的な文、例えば、音声を利用した情報告知や電話を利用した音声案内サービスに利用されている文に対応する音声合成では、任意の文章ほど複雑でないので、予め文の構造やパターンに対応した韻律データをデータベースとして蓄積しておき、韻律パラメータを計算する場合には、蓄積されたパターンを検索し類似のパターンの韻律パラメータを利用することができる。この方法を用いることにより、テキスト音声合成方法により得られた合成音と比較して、自然性を著しく改善できる。例えば、特開平11-249677号公報には当該方式を利用した韻律パラメータ計算方法が開示されている。

【0007】

合成音声の抑揚やイントネーションは韻律パラメータの品質に依存する。また、適切に制御することにより、感情表現や方言等の合成音の話調を制御することが可能である。

【0008】

これらの定型的な文の関する従来の音声合成技術は、主として音声を利用した情報告知や電話を利用した音声案内サービスに利用されているが、その利用形態

においては、合成音声は1つの話調に固定され、方言や外国語音声等多様な音声を任意に合成することが不可能であった。方言等は携帯電話や玩具等、アミューズメント性を必要とする装置への搭載が望まれ、また外国語音声に関しては、国際化には必須の技術である。

【0009】

【発明が解決しようとする課題】

しかし、従来の技術では、音声合成時に、各方言や言い回しに発声内容を随意変換することは考慮されておらず、技術上困難であり、システム利用者及び運用者以外の第三者が自由に上記韻律データを作成することは困難であった。更に、携帯電話端末のような計算用資源が極度に限定されており、音声合成プログラムの変更が困難である装置において、上述の多様な話調の音声を合成することができなかった。

【0010】

発明の主な目的は、音声合成手段が搭載されている端末内で定型的文に対する多種の話調の音声を合成するための音声合成方法及び装置を実現することである。

【0011】

発明の他の目的は、音声合成装置の製造者、所有者、利用者以外の第三者が「韻律データ」を作成し、音声合成装置の使用者がそのデータを利用できる韻律データ配信方法を提供することである。

【0012】

【課題を解決するための手段】

上記目的を達成するため、本発明の音声合成方法では、合成音声により出力すべき発声内容種別を特定する複数の内容識別子を設け、それぞれの内容識別子に対して複数種の話調の韻律データが格納された話調辞書を作成し、音声合成の実行時に、上記内容識別子及び上記話調を指定することにより、上記話調辞書から指定された韻律データを読み出し、読み出された韻律データを音声合成駆動データとして音声に変換する。

【0013】

また、本発明による音声合成装置は、合成音声により出力すべき発声内容の種別を特定する内容種別を識別する識別子発生する手段と、上記合成音声により出力すべき発声内容の話調を指定する話調指定手段と、複数の内容識別子のそれぞれに対応する複数の話調及び上記内容識別子及び話調に対応付けられた韻律データからなる話調辞書と、上記内容識別子及び話調が指定されたとき上記話調辞書から上記指定された内容識別子及び話調の韻律データを読み出し音声に変換する音声合成処理部とをもつ。

【0014】

上記話調辞書の作成は、音声合成装置又は音声合成装置をもつ携帯端末等の製造時に前もって、音声合成装置又は端末に組み込む他に、通信ネットワークを介して、必要な内容識別子及び任意の話調の韻律データのみを取り込む、或いは移動可能な小型メモリにして、端末で着脱できるようにしても良い。話調辞書の作成は、発声内容管理方法を端末の製造者、ネットワークの管理者以外の第三者に開示し、その発声内容管理方法に従って、内容識別子と対応する韻律パラメータからなる話調辞書を作成させてもよい。

【0015】

本発明により、音声合成装置或いは音声合成装置を備える端末に組み込むプログラムの開発者は、合成すべき話調を指定する話調指定子と内容識別子のみの情報から、所望話調の音声合成を実現できる。また、話調辞書作成者は、合成プログラムの動作を考慮に入れることなく、文識別子に対応する話調辞書を作成するだけでよいので、簡便に所望の話調での音声合成を実現できる。

【0016】

【発明の実施の形態】

図1は、本発明による音声合成装置及び音声合成方法が実施される情報配信システムの一実施形態を示すブロック図である。

【0017】

本実施形態の情報配信システムは、本発明による音声合成装置をもつ携帯電話機等の端末装置（以下単に端末と呼ぶ）7が接続可能な通信網（ネットワーク）3と、通信網3に接続された話調辞書格納サーバ1、4とを有し、端末7は、端

末使用者 8 が指定した話調に対応する話調辞書を指定する手段と、指定された話調辞書をサーバ 1、4 から端末に転送するデータ転送手段と、転送された話調辞書を、端末 7 内の話調辞書格納メモリに格納する話調辞書格納手段を備えることにより、端末使用者 8 が希望する話調で定型文的な合成音を出力する。

【 0 0 1 8 】

携帯端末使用者 8 が上記話調辞書を利用して合成音の話調を設定する形態について説明する。

【 0 0 1 9 】

第一の方法は、製造者等の端末供給者 9 が端末 7 に話調辞書を搭載するプレインストール方法である。この場合は、データ作成者 10 が話調辞書を作成し、それを携帯端末供給者 9 に提供し、携帯端末供給者 9 は話調辞書を携帯端末 7 のメモリに格納し、携帯端末 7 を携帯端末使用者 8 に供給する。この第一の方法では、携帯端末使用者 8 は、携帯端末 7 の使用開始時から出力音声の話調の設定、変更ができる。

【 0 0 2 0 】

第二の方法は、データ作成者 5 は、携帯端末 7 が接続可能な通信網 3 を所有する通信事業者 2 に対し話調辞書を供給し、通信事業者 2 ないしはデータ作成者 5 が話調辞書格納サーバ 1、4 に話調辞書を格納する。通信事業者 2 は、携帯端末使用者 8 から端末 7 を通じて話調辞書の転送要求（ダウンロード）を受けると、話調辞書格納サーバ 1 に格納されている話調辞書を携帯端末 7 が取得可能かどうかの判定を行う。この際、話調辞書の特質に応じて通信料もしくは取得量を携帯端末使用者 8 に請求してもよい。

【 0 0 2 1 】

第三の方法は、端末使用者 8、端末製造者 9、通信事業者 3 以外の第三者 5 が話調辞書を作成し、第三者のデータ作成者 5 は、発声内容管理リスト（定型的文の種別を表す識別子の対応データ）を参照し、話調辞書を作成し、話調辞書格納サーバ 4 に話調辞書を格納する。話調辞書格納サーバ 4 は、通信網 3 を通じて端末 7 からアクセスされ、端末使用者 8 の要求に応じて話調辞書の取得を許可する。その話調辞書を取り込んだ端末 7 の所有者 8 が所望の話調を選択して端末 7 か

ら出力する合成音声メッセージ（定型的文）の話調を設定する。この際、データ作成者 5 は話調辞書の特質に応じたライセンス料を、通信事業者 2 を代行者として携帯端末使用者 8 に請求してもよい。

上記 3 つの何れかの方法を用いて、端末使用者 8 は、携帯端末 7 において出力される合成音声の話調を設定、変更するための話調辞書を取得する。

【0022】

図 2 は、本発明による音声合成装置をもつ端末である携帯電話機の一実施形態の構成を示す図である。携帯電話機 7 は、アンテナ 18、無線処理部 19、ベースバンド信号処理部 21、入出力部（入力キー、表示部など）及び音声合成装置 20 をもつ。音声合成装置 20 以外の部分は従来知られているものと同じであるので説明を省く。

【0023】

同図において音声合成装置 20 の話調辞書指定手段 11 は、端末 7 の外部から話調辞書を取り込むときに、発声内容識別子入力手段 12 で指定された内容識別子を使用して話調辞書を取り込むものである。発声内容識別子入力手段 12 は発声内容識別子を入力するもので、例えば、携帯端末 7 がメールを受信した時に、自動的に識別子がメール受信報知メッセージであることを表す識別子をベースバンド処理部 21 から入力する。

【0024】

話調辞書格納メモリ 14 は、その詳細は後述するように、発声内容識別子に対応する話調及び韻律データを記憶するメモリで、データはプレインストールされる場合と、通信網 3 を介してダウンロードされる場合がある。韻律パラメータ格納メモリ 15 は、話調辞書格納メモリ 14 から選択された特定の語調の合成音のデータを格納するメモリである。合成波形格納メモリ 16 は話調辞書格納メモリ 14 のデータを波形信号に変換して記憶するメモリである。音声出力部 17 は合成波形格納メモリ 16 から読み出された波形信号を音響信号として出力するもので、電話機のスピーカと兼用される。

【0025】

CPU 13 は上記各手段、メモリを駆動、制御し音声合成を行うためのプログ

ラムが格納されている信号処理装置で、ベースバンド処理部21の他の通話処理のための処理を行うCPUと共用してもよい。説明の都合上音声合成部の構成素子として示されている。

【0026】

図3は上記発声内容識別子を説明する図で、複数の識別子とそれぞれの識別子の表す発声内容との対応リストを構成している。同図では、識別子「ID_1」、「ID_2」、「ID_3」及び「ID_4」に対しては、それぞれの識別子に対応する発声内容の種別「メール着信報知メッセージ」、「通話者着信報知メッセージ」、「発信者報知メッセージ」及び「アラーム情報報知メッセージ」が定義されている。

【0027】

話調辞書作成者5又は10は、例えば「ID_4」という識別子に対して、「アラーム情報報知メッセージ」であるところの任意の話調辞書を作成できる。なお、図3の関係は秘匿すべきものではなく、書類（音声内容管理データテーブル）として広く公開する。もちろん、電子的データとして計算機上及びネットワーク上で公開してもよい。

【0028】

図4及び図5はいずれも上記識別子に対し、話調の異なった例として、標準語と大阪方言の発声内容文を示す。図4は話調が標準語の発声文（以下、「標準パターン」と表記）を示す。図5は話調が大阪方言の発声文（以下、「大阪方言」と表記）を示す。例えば、識別子「ID_1」に対しては、標準パターンでは、「メールを着信しました」という発声文内容として、大阪方言においては、「メールが来てまっせ」という発声文内容を記述する。これらの文言は、話調辞書を作成する作成者が任意に定義できるものであり、上記例とする必要はない。例えば、大阪方言の識別子「ID_1」に対しては、「来ました、来ました、メールでっせ！」でも良い。また、図5の識別子「ID_4」のように、文の一部（○で示す文字）を入れ替えることのできる定型文でもよい。

【0029】

このようなデータは、発信者情報のように固定的に準備できない情報を読み上

げるのに有効である。定型的な文を読み上げる方法は、文献「単語及び文韻律データベースを用いた韻律制御方式の検討」（日本音響学会講演論文集、pp. 227-228、1998）に開示されている技術が使用できる。

【0030】

図6は、上記話調辞書の1実施形態におけるデータ構造を示す図である。このデータ構造は、図2の話調辞書格納メモリ14に格納される。話調辞書は、いずれの話調であるかを表す話調識別情報402、インデックステーブル403、各識別子に対応する韻律データ404～407から構成される。話調識別情報402は、話調辞書14の話調の種別を示す、例えば、「標準パターン」「大阪方言」等の種別を登録する。また、話調辞書14に特徴的なシステム内共通の識別子を付与しても良い。話調識別情報402は、端末機7において、話調を選択する際のキー情報となる。インデックステーブル403は、各識別子に対応する話調辞書の始まる先頭番地を示すデータが格納される。端末機において識別子に対応する話調辞書を探索する必要があり、インデックステーブル403により管理することで、高速の検索ができる。もちろん、各韻律データ404～407を固定長のデータとし、順次探索するような方法を採用すれば、インデックステーブル403を設ける必要はない。

【0031】

図7は、図6に示した各識別子に対応する韻律データ404から407のデータ構造を示す。図2の韻律パラメータ格納メモリ15に格納される。韻律データ501は、識別子502及び音素テーブル503から構成される。識別情報子502には、韻律データの発声内容識別子を記述する。例えば、図4の「ID_4」と「〇〇の時間になりました」の例であれば、「ID_4」と記述する。一方音素テーブル503は、音声合成装置駆動データ、すなわち、発声文内容の音素表記、各音素の長さ、各音素の高さからなる韻律データである。ここで、一例として、大阪方言の話調辞書における識別子「ID_1」に対応する発声内容である「メールが来てまっせ」に対する音素テーブルを図8に示す。音素テーブル601は、音素表記602、音素の長さ603、音素の高さ604のデータで構成される。音素の長さはミリ秒単位で示されているが、音素の長さを表記できる物

理量であれば、この限りではない。同様に、音素の高さはヘルツ単位で示されているが、高さを表現できる物理量であれば、この限りでない。

【 0 0 3 2 】

本例では、音素の表記は図 8 に示すとおり、「m / e / e / r / u / g / a / k / i / t / e / m / a / Q / s / e」となる。また、音素「r」に対応する音素の長さは 3 9 ミリ秒であり、高さは 3 5 2 ヘルツであることを示している（6 0 5）。表記中「Q」6 0 6 は促音を意味する音素記号である。

【 0 0 3 3 】

図 9 は、本発明による音声合成方法の一実施形態における話調の選択から合成音声波形を生成するまでの生成手順を示す。ここでは、一例として、図 2 の携帯端末 7 の使用者が「大阪弁」の合成話調を選択し、通話着信時に合成音によるメッセージを流す実施方法を示す。管理テーブル 1 0 0 7 は、通話着信時に合成内容を決定するために用いるための電話番号及び人名情報を格納する。

【 0 0 3 4 】

上記例に対して波形を合成する場合、まず、話調辞書指定手段 1 1 から入力された話調辞書指定情報により、話調辞書格納メモリ 1 4 の話調辞書を切り替える（S 1）。話調辞書格納メモリ 1 4 に話調辞書 1（1 4 1）又は話調辞書 2（1 4 2）を格納する。携帯端末 7 の通話着信時には、発声内容識別子入力手段 1 2 において、識別子「ID__2」を用いて「通話着信報知メッセージ」を合成する旨を決定し、識別子「ID__2」を合成対象の韻律データとする（S 2）。続いて、発生すべき韻律データを決定する（S 3）。本例の場合、任意に語彙を入れ替える文ではないので特に処理は行わない。しかし、例えば、第 5 図の「ID__3」の発声内容を利用する場合には、管理テーブル 1 0 0 7（図 2 のベースバンド処理部 2 1 にも受けられている。）より、発呼者の人名情報を取得し、「すずきさんからやでえ」という韻律データを決定する。

【 0 0 3 5 】

以上のようにして韻律データを決定した後、図 8 で示される音素テーブルを計算する（S 4）。上記例の「ID__2」を利用して合成する場合、話調辞書格納メモリ 1 4 に格納されている韻律データを韻律パラメータ格納メモリ 1 5 に転送

するだけで良い。

【0036】

しかし、例えば、第5図の「ID_3」の発声内容を利用する場合には、管理テーブル1007より、発呼者の人名情報を取得し、「すずきさんからやでえ」という韻律データを決定する。「すずき」の部分の韻律パラメータを計算し、韻律パラメータ格納メモリ15に転送する。「すずき」の部分の韻律パラメータの計算は、例えば、文献「単語及び文韻律データベースを用いた韻律制御方式の検討」（日本音響学会講演論文集、pp. 227-228、1998）に開示されている方法を利用することができる。

【0037】

最後に、CPU13が、韻律パラメータ格納メモリ15に格納された韻律パラメータを読み出し、それに対応した合成波形データに変換し合成波形格納メモリ16に格納する（S5）。合成波形格納メモリ16の合成波形データは順次音声発生部すなわち電気・音響変換機17によって、合成音声として出力される。

【0038】

図10及び図11は、いずれも本発明による音声合成装置を備えた携帯端末で、合成音声の話調を指定する際の端末の表示画面を示す図である。端末使用者8が、携帯端末7の表示画面71で「合成話調設定」メニューを選択する。図(a)では、「合成話調設定」71aは、「アラーム設定」や「着信音設定」と同一階層で実現されているが、同一階層である必要はなく、合成話調設定の機能が実現されていれば他の方法でもよい。合成話調設定メニュー71aが選択された後には、図(b)のように携帯端末7に登録されている合成話調を表示画面71に表示する。ここで表示されている文字列は、図6の話調識別情報402に格納されている文字列である。例えば、話調辞書がねずみが話す様態の音声を出力させるために作成されたデータである場合、「ネズミでちゅー」というような文字列を表示する。もちろん、上記話調辞書の特徴が示される文字列であれば、他の表記文字列であってもよい。例えば、携帯端末使用者8が「大阪弁」で合成させたいとの意思を持っている場合、「大阪弁」の表示71bを反転させ、合成話調を選択する。また、話調辞書には、日本語だけでなく、「英語」「フランス語」の話調辞書

もしくは発音表記で格納してもよい。

【0039】

図11は、図1の携帯端末使用者8が通信網3を経由して話調辞書を取得する方法を説明するための携帯端末の表示部を示す図である。携帯端末7は通信網3を経由して情報管理サーバに接続した際表示される画面であり、(a)は本発明の話調辞書配信サービスに接続した後の画面である。

【0040】

まず、携帯端末使用者8に対して、合成話調データを取得するかどうかを確認する画面71を表示し、了解を意味する「OK」71cを選択した場合には、画面71を(b)に切り替え、情報管理サーバに登録されている話調辞書の一覧を表示する。ここでは、ねずみの模倣音声である「ネズミでちゅー」、大阪弁口調のメッセージである「大阪弁」等の話調辞書が登録されている。

【0041】

次に、携帯端末使用者8は取得したい話調データに反転表示を移動させ、確認ボタンを押下する。情報管理サーバ1、では、要求された話調に対応する話調辞書を通信網3に送出する。送出が完了した後、話調辞書の送受信を完了する。以上の手順で携帯端末7に存在しない話調辞書を携帯端7内に格納する。上述の方法では、通信事業者の提供するサーバにアクセスしデータを取得したが、もちろん、通信事業者ではない第三者5が提供する話調辞書格納サーバ4にアクセスしてデータを取得する方法でもよい。

【0042】

【発明の効果】

本発明により、定型的な情報の読み上げを、任意の話調で読み上げることが可能な携帯端末を簡便に開発することができる。

【図面の簡単な説明】

【図1】

本発明による音声合成装置及び音声合成方法が実施される情報配信システムの一実形態を示すブロック図である。

【図2】

本発明による音声合成装置をもつ端末である携帯電話機の一実施形態の構成を示す図である。

【図 3】

発声内容識別子を説明する図である。

【図 4】

標準語識別子に対する発声内容文を示す図である。

【図 5】

大阪方言の識別子に対する発声内容文を示す図である。

【図 6】

話調辞書の一実施形態におけるデータ構造を示す図である。

【図 7】

図 6 に示した各識別子に対応する韻律データのデータ構造を示す図である。

【図 8】

図 5 の話調辞書における大阪方言「メールが来てまっせ」に対する音素テーブルを示す図である。

【図 9】

本発明による音声合成方法の一実施形態の音声合成手順を示す図である。

【図 1 0】

本発明による携帯電話機の一実施形態における表示部を示す図である。

【図 1 1】

本発明による携帯電話機の一実施形態における表示部を示す図である。

【符号の説明】

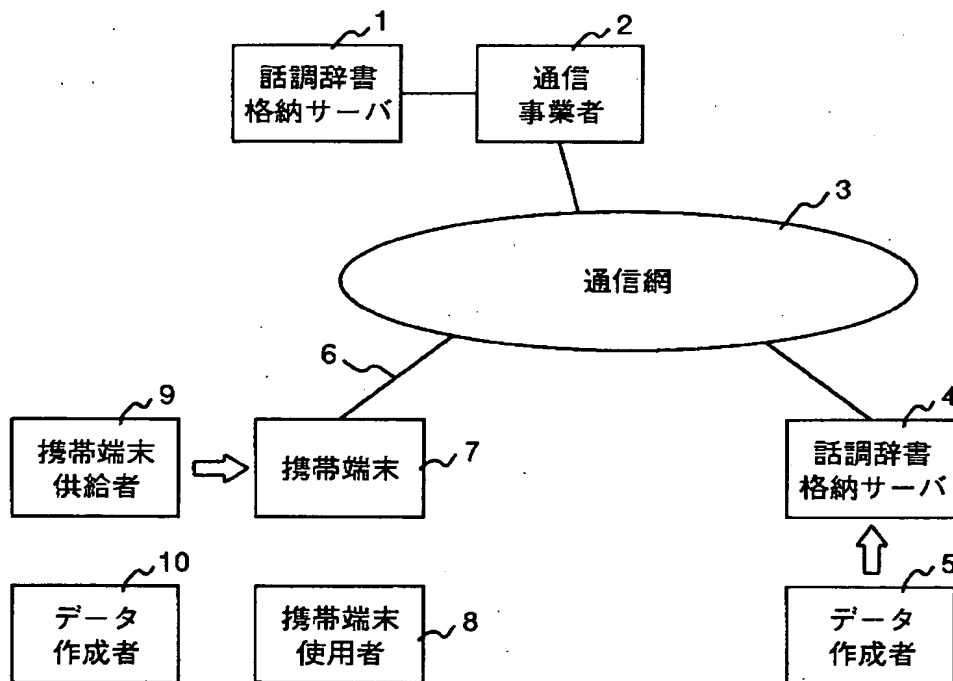
1 : 話調辞書格納サーバ、 2 : 通信事業者、 3 : 通信網、 4 : 話調辞書格納サーバ、 5 : データ作成者、 6 : 通信回線、 7 : 携帯端末、
8 : 携帯端末使用者、 9 : 携帯端末供給者、 1 0 : データ作成者、
1 1 : 話調辞書指定手段、 1 2 : 発声内容識別子入力手段、
1 3 : 音声合成手段、 1 4 : 話調辞書格納メモリ、
1 5 : 韻律パラメータ格納メモリ、 1 6 : 合成波形格納メモリ、
1 7 : スピーカ、 1 8 : アンテナ、 1 9 無線処理部、 2 0 : 音声合成装置

、 2 1 : ベースバンド信号処理部、 2 2 : 入出力部、
4 0 1 : 話調辞書データ構造例、 4 0 2 : 識別情報、
4 0 3 : インデックステーブル、 4 0 4 : 韻律データ、
4 0 5 : 韻律データ、 4 0 6 : 韻律データ、 4 0 7 : 韻律データ
5 0 1 : 韻律データ構造例、 5 0 2 : 識別情報、 5 0 3 : 音素テーブル、
6 0 1 : 音素テーブル例、 6 0 2 : 音素表記項、 6 0 3 : 長さ項、
6 0 4 : 高さ項、 6 0 5 : 音素表記「r」に対する韻律パラメータ例
6 0 6 : 促音表記「Q」、 7 1 : 表示画面、 S 1 : 合成話調選択ステップ
S 2 : 合成内容決定ステップ、 S 3 : 韻律データ決定ステップ、
S 4 : 韻律パラメータ計算ステップ、 S 5 : 波形合成ステップ、
1 0 0 7 : 管理テーブル。

【書類名】 図面

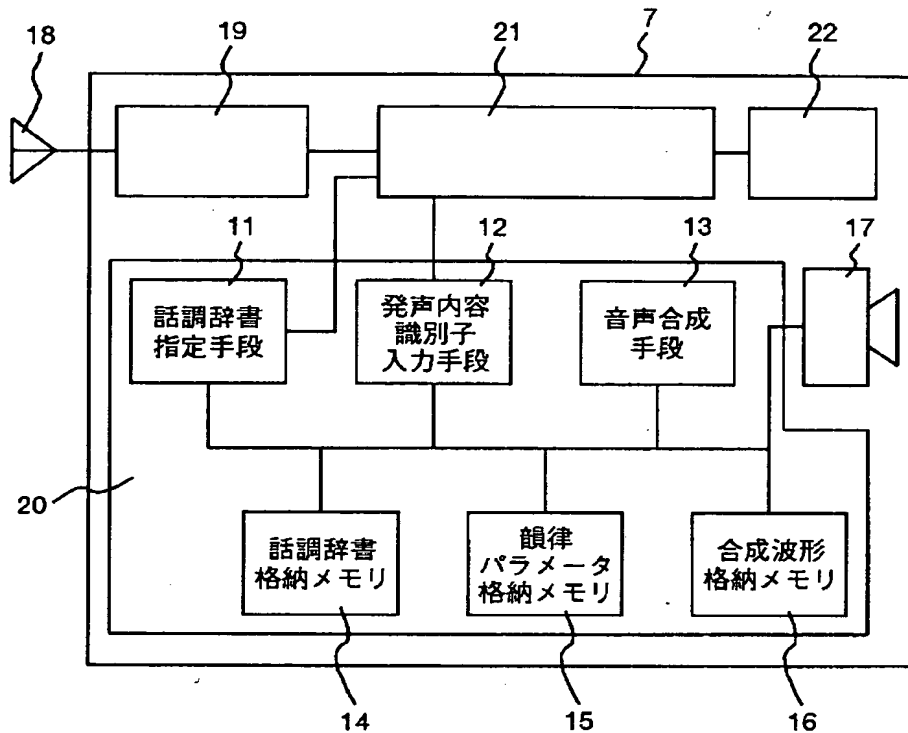
【図 1】

図 1



【図 2】

図 2



【図 3】

図 3

識別子	識別子内容
ID_1	メール着信報知メッセージ
ID_2	通話着信報知メッセージ
ID_3	発信者報知メッセージ
ID_4	アラーム情報報知メッセージ

【図 4】

図 4

識別子	発声内容文
ID_1	メールを着信しました
ID_2	電話がかかっております
ID_3	〇〇さんから電話がかかっております
ID_4	〇〇の時間になりました

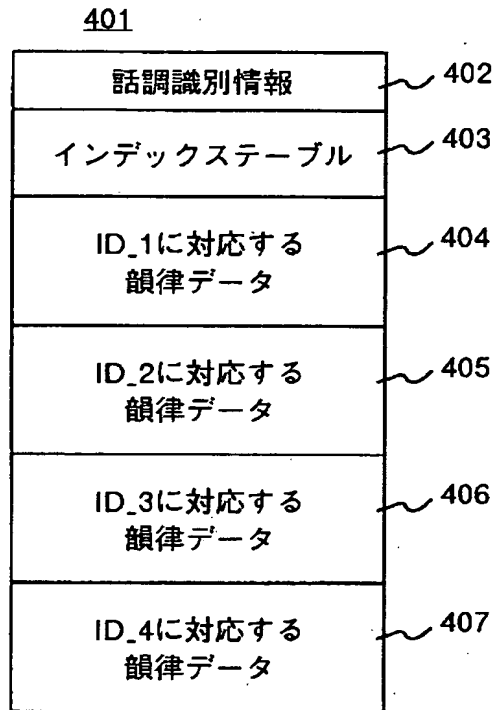
【図 5】

図 5

識別子	発声内容
ID_1	メールが来てまっせ
ID_2	電話でっせ
ID_3	〇〇さんからやでえ
ID_4	〇〇の時間になったんやけど

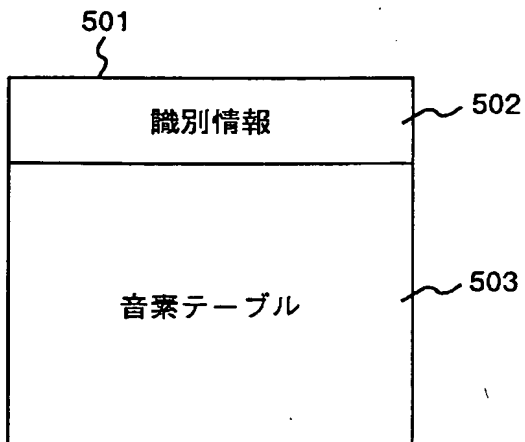
【図 6】

図 6



【図 7】

図 7



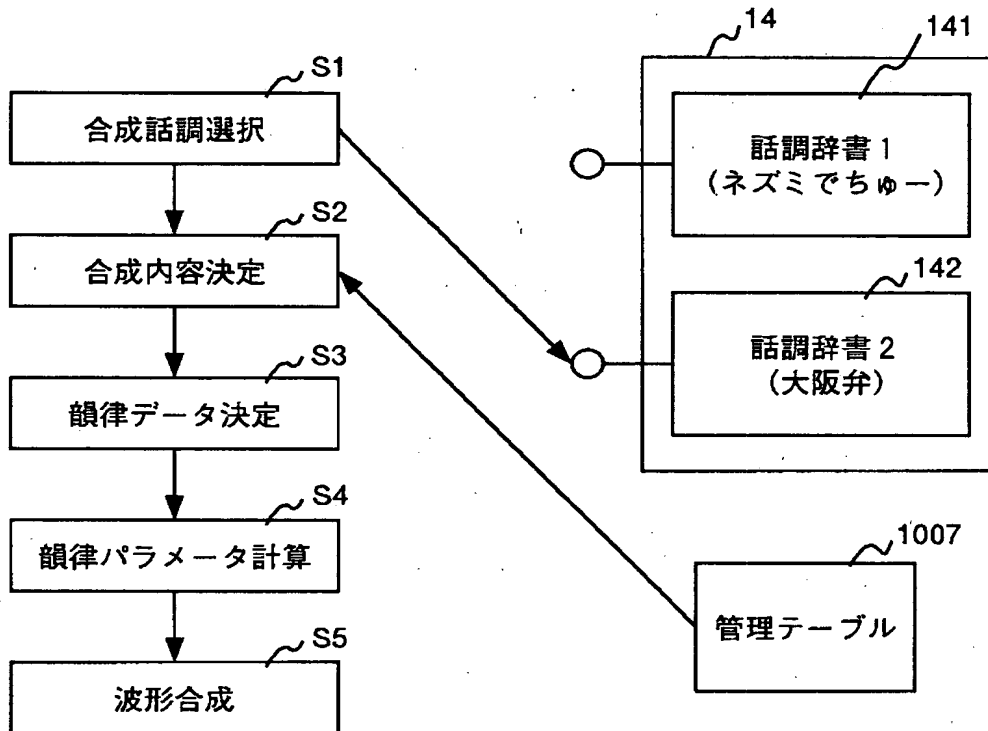
【図 8】

図 8

音素表記	長さ	高さ
m	29	247
e	72	253
e	116	337
r	39	352
u	79	356
g	40	353
a	88	343
k	77	337
i	54	352
t	53	359
e	75	328
m	55	299
a	56	274
Q	125	234
s	78	214
e	65	202

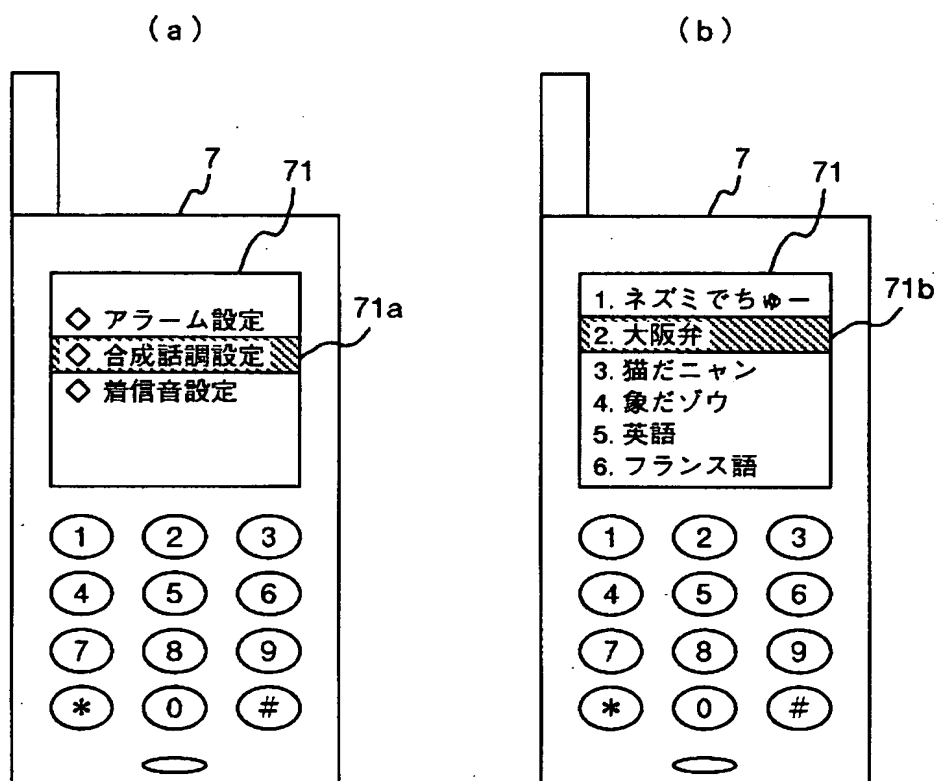
【図 9】

図 9



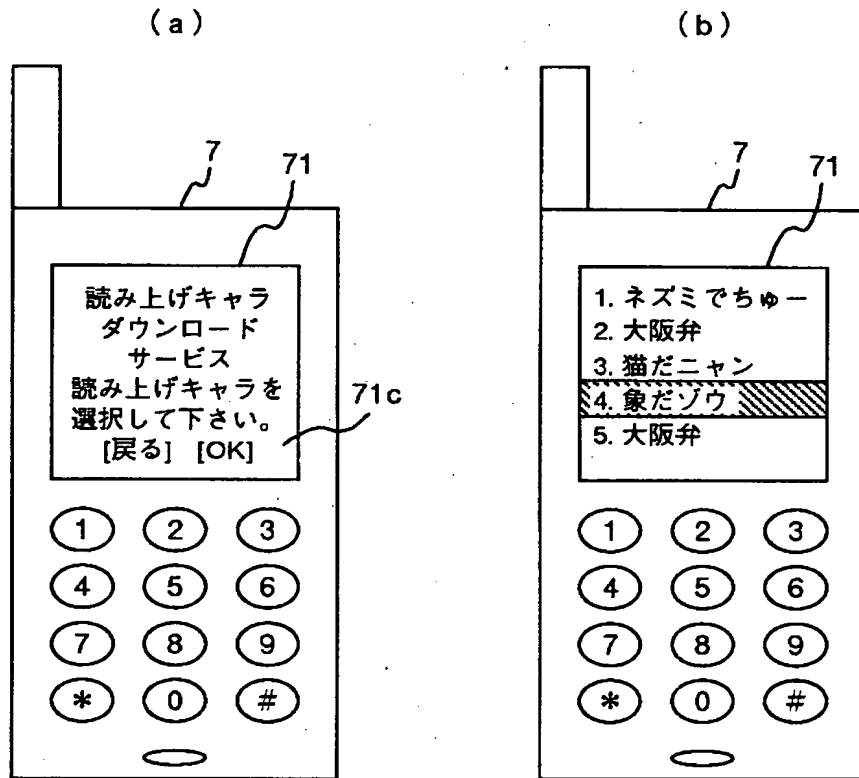
【図10】

図 10



【図 11】

図 11



【書類名】 要約書

【要約】

【課題】 定型的文を任意の話調の音声に音声合成する。また、第三者が韻律データを作成し、音声合成部をもつ端末の利用者がその韻律データを取り込み利用できる方法を提供する。

【解決手段】 定型的文の発声内容の種別を指定する発声内容識別子定め、上記内容識別子に対応する話調及び韻律データからなる話調辞書 1 4 を作成し、発生すべき合成音声の内容識別子及び話調を指定（1 2）して、話調辞書 1 4 から発生すべき合成音声の韻律データを選択し（1 5）、選択された韻律データを音声合成装置駆動データとして音声合成装置 1 3 に加え、特定の話調の音声合成を行う。

【選択図】 図 2

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地

氏 名 株式会社日立製作所